Why people minimax under ambiguity

1. Perhaps they weight the likelihoods of the worst money payoffs as a way of reacting to risk, i.e., instead of "worsening" the payoff in terms of utilities. This seems to be what Fellner is suggesting in Competition Among the Few; people do this because the worst outcomes have unfortunate side effects.

Then why would they weight the worst even more under ambiguity? (This is implied, if they use minimax procedures even when von Neumann-Morgenstern utilities are involved--if they do). Maybe Fellner's suggestion in his new paper applies: A loss suffered in an ambiguous situation--as opposed to a "certainty," or the option of not betting--may be "less acceptable" than a gain in that ambiguous situation, weighted by the "best guess" estimate of the gain's likelihood. In other words, a certain loss may be acceptable in prospect if the subject can tell himself, I had a definite x% chance of so much gain; but he may feel more of a fool in the ambiguous situation exposing himself to this regret, unless the estimated % is even higher, or unless the possible prize is higher for the same estimated probability.

2. What is the relation between von Neumann-Morgenstern utilities and "variance preferences"? How is a dislike for high variance in terms of money translated into low expectation in terms of utilities; what is true of the variance of the new bet in terms of utilities?

Was it ever reasonable to suppose that one avoided variance in terms of money? Wasn't it always "downward variance," and could this have been expressed by minimaxing distributions?

Do the people following the Hodges and Lehmann formula favor actions whose expectations are insensitive to the ambiguity; or are the variances also

4. Preferences for (against) ambiguity are distinct from preferences for or against risk, as the latter reveal themselves in terms of preferences for variance or skewness. I conjecture that (a) between two bets with equal expectations, variance, skewness in terms of money, individual will prefer the one whose probability distribution is "better known," "less ambiguous", or vice versa. (b) same in terms of utility payoffs (?)

Thus, individual may have a liking or dislike for "ambiguity" which is independent of (though possibly related to) his like or dislike for risk; he might conceivably like "high-risk" gambles (with given expectation of money or utility) yet dislike "ambiguity": prefer a known-risk gamble with lower variance or even lower variance and expectation (or at least, with same variance and expectation) to an ambiguous gamble.

Experiments so far have only tested preferences under risk, even when nominally testing the Savage axioms. (Incidentally: what is the decision rule underlying people's variance preferences and skewness preferences, in Coombs' examples? Can one be constructed that is consistent with the axioms?)

5. Thus, as Keynes says, "weight of the argument" isn't equivalent to "probable error," since that refers to a given probability distribution (it might, instead, have something to do with the range of the set of distributions among which it is difficult to discriminate); new evidence might lead to a definite, confidently-held probability distribution which had a greater "probable error" than the old, ambiguous one.

On the other hand, new evidence doesn't invariably add to the "weight of the argument," either, as Keynes implies; at least, it may not decrease ambiguity. If it conflict with something said earlier, or if it casts doubt on the reliability of earlier evidence, it may increase ambiguity.

6. Consider properties of "status quo" strategy; typically, regarded as having less ambiguous outcome; even, constant outcome. Thus,

7. Perhaps Ellsberg rule, or something like it, is answer to the "group decision problem," with the set of reasonable distributions being the set advanced by one or another member. (This is a reasonable metaphor even for single-person decision under ambiguity).

8. Knight may be talking mainly about situation in which subjective probabilities differ; this is likely to be low information situation, in which it is also true that many people do not obey the axioms, i.e., do not assign probabilities, but main phenomena he is interested in may reflect merely the fact of difference in probabilities. We are talking about the same situation; Savage emphasizes that probabilities can always differ, but it is in low information situation that they are most likely to differ, and to differ by most.

thus: WHAT PART OF KNIGHTIAN BEHAVIOR IS TO BE ACCOUNTED FOR BY DIFFERING PROBABILITIES (which would still be distinctive behavior worthy of being identified, as associated with low information situations) AND WHAT PART IS DETERMINED BY NON-PROBABILITISTIC DECISION RULES?

9. Measure of "uncertainty": - sum p_x log p_x refers only to a given probability distribution; says nothing about "confidence" (or "value of information"). Like "probable error." (?)

10. Degree of ignorance or ambiguity is related to our expectation that an additional piece of evidence might greatly, or "significantly," or critically, change our overall probability distribution. (e.g., we might assume that even a small sample, or a single observation, would greatly affect our "a priori" probability distribution, which might be rectangular, or asymmetric but of low confidence.

11. Low information states, or ignorant states, are characterized behaviorally and subjectively not only by a) differences in estimates; b) low confidence (sometimes) in estimates; (c) high "probability of error" or "variance" or "uncertainty" (above) in individual estimate of prob distribution; but by (Klein) TENDENCY TO ENGAGE IN INFORMATION-SEEKING, SEARCHING, EXPERIMENTAL BEHAVIOR, and high value to sequential-decisionmaking. LEARNING STRATEGIES AND BEHAVIOR.

12. Use nuclear weapons tests as example of analysis: value of information.

13. If a sequence of decisions is foreseen such that learning could take place--uncertainty could be reduced, subjective probabilities changed ((two ways to characterize learning process))--then "Ellsberg behavior" might characterize the early stages of that process; or some other, "search" behavior might do it. But maybe the payoffs associated with "search" actions

in

13 continued:

in the early stages of what is perceived as a learning process have just the characteristics thqt my decision rule would favor. That is, maybe my decision rule selects "experimental," sampling, observational, searching, behavior in conditions of ignorance.

14. A difference between "ignorance" and "risk" is the difference in attitudes toward the "higher rank uncertainties" (see Fisher). In Urn I, does person act toward all the various compositions of the urn as if they were eqqally likely? (i.e., obey axioms w.r.t. bets on the composition of the urn?)) I suspect not.

15. e.g.,: in Urn I we have Red, Yellow or Black balls, unknown ratio; in Urn II we have 30, 30, 30.
a) Which would you rather bet on, (Yellow or Black)$_{II}$ or (Yellow or Black)$_{I}$, or are you indifferent?

b) Which would you rather bet on, Red$_{I}$ or Red$_{II}$?

((if first answer is (Yellow or Black)$_{II}$ and second is Red$_{II}$, contradiction; if second answer is indifference, contradiction.
If second answer is Red$_{I}$, no contradiction yet, but there will be.))
c) (If second answer has been Red$_{I}$): Which would you rather bet on, Red$_{I}$, Yellow$_{I}$, Black$_{I}$? Which would rather bet on, R_{II}, Y_{II}, B_{II}?
OR (d) Which would you rather bet on, RI or YII? BI or BII?

Imagine sample of 3: R, Y, B. Or sample of 2: Red + Yellow.

16. Example of decision that could have been sequential, but wasn't: Decision to drop two atom bombs on Japan, made before any bombs have been tested (decision not reconsidered after uncertainty as to whether bomb would work was substantially reduced--at least, for scientists. Note: test didn't reduce uncertainty for politicians nearly as much; effect of test on reducing ignorance or uncertainty generally different for expert and for layman, with layman not necessarily always wrong.)

17. Example of difficulty of making a decision process truly sequential: importance of sunk costs, difficulty of stopping process or cutting costs on basis of new information: Offensive in Flanders (see In Flanders Fields).

18. Certain types of behavior in response to behavior presume consciousness of uncertainty: deliberate conservatism, search behavior. Others might arise even in connection with subjective certainty, in circumstances that could "objectively" be characterized as "low information": situations where estimates differ, change significantly with small increases in information, prior experience scanty or irrelevant or conflicting. (e.g., unconscious wishful thinking possible effect); there are times when one "should" be uncertain, or act "as if" you were uncertain, even when you feel certain. ((e.g., have possibility of STOP message to SAC planes)). Group decision problem.

19: (on Klein): Multiples paths could just be a reaction to a given, high variance probability distribution: to raise expected value of outcome. But Klein emphasizes that it is especially, or also, worthwhile when there is a chance to change the probability distribution or reduce its variance (or, more precisely, to gain "valuable information") during the course of development; this might justify multiple paths even when initial probabilities did not have high variance. ((In fact, this may be precisely a case where "subjective certainty at start of project is not to be trusted; you should act as if you were uncertainty even if you feel certain."))
Likelihood of surprise; unanticipated events; major changes in probability distribution.

20. Maybe definition of "surprise" is: event that significantly changes our subjective probability distribution. "Surprises" in this sense are "likely," or "to be expected" in situation of ignorance, which is likely to be "start of a learning process" (if learning is possible); they are less likely, and less anticipated, at later stage in learning process, when we think we know likelihoods even though outcomes may be stochastic; hence, more "surprising" later. (One significant change would be putting some positive probability on events earlier considered "impossible," or zero possibility.)

21. Clausewitz' notes on info show clearly how early combat information may lead to high confidence, low subjective ambiguity, on false basis; with new information either leading to totally different picture (no more certain, no more true), or to great increase in subjective ambiguity.

22. Clausewitz also shows (p. 51) how earlier information may be more accurate ~~[illegible]~~ (pre-hostilities calculations, more deliberate, less prone to panic or small sample effects) than somewhat later ~~[illegible]~~ (initial info in combat situation, which gets weight disproporionate to size of sample--inevitably).

23. The Importance of Being Uncertain. Important to know when you "should" act "as if" you were uncertain. And how is that? ((Unless decision rule under perceived uncertainty is so conservative that you would never take ~~any~~ any positive action under it; always choose status quo.))z ((Don't want most pessimistic assumptions to dominate decision. Is that the curse of "committee decision," when that <u>is</u> a curse?)) (May mean, Acting as though a committee decision rule were necessary, where members of committee have ~~different~~ different subjective probabilities.)) ((Latter case may indicate a situation of ignorance, in which there is a fair chance that <u>none</u> of them are "right").

24. There has always been recognition of the intuitive difference between ignorance and risk: but denial that one "should", or does, reflect this difference in one's behavior. Now, two types of behavior may reflect this difference: a) (me): under ignorance, individual may follow decision rule that violates Savage axioms; (b) (Klein): learning or search strategies may be more useful, and more employed (e.g., by "successful" R&D firms) under ignorance than under risk. In latter case, I would like to be able to answer, "How <u>many</u> multiple paths should be pursued, given costs? How much is experimental information worth?" May come from relating (b) to (a); show how "learning actions" look in terms of decision rule for ignorance.

25. The WORTH OF AN EXPERIMENT may be related to, [illegible]
a) the difference in expected payoff before the experiment and after the experiment ((BAD: this could make the experiment look very worthwhile even though it didn't lead to different action; or exaggerate worth if it did));
b) difference in expected payoff to "best action" before the experiment and after the experiment, both computed in terms of probabilities after the experiment (better; but this requires knowledge of probabilities after the experiment, though before consequences of action are known).
(c) expected difference in [illegible] expected payoffs under (b); (this assumes a known probability distribution of the distributions that might result from the experiment). (Marschak solution)
(d) some function of the set of expectations (corresponding to different "best actions") corresponding to different possible distributions resulting from the experiment; e.g., range, minimum, maximum, "average" (on arbitrary equal weighting assumption).

e) The "best act" prior to experiment need not be one with highest expected value (even in terms of von Neumann-Morgenstern payoffs); may be minimax, or "Ellsberg choice," or minimax regret; best act after experiment must be compared with this.

We must consider, in evaluating worth of experiment: Should our measure be capable of evaluation prior to experiment? Or may it reflect results of experiment? Or might it reflect actual consequences of action?

(Experiment--and messages, or "information" in general--may not always "decrease uncertainty" (e.g., in sense of variance of subjective prob dist); it may increase risk, or increase ambiguity, or both; or it may leave "degree of uncertainty" unchanged but shift distribution (perhaps both old and new positions being, in fact, "wrong.")

26. If parameter "confidence," or "credibility," etc., were significant only when relative likelihoods were equivalent--as Georgescu-Roegen (and Chipman) imply--it would be of trivial importance. (It would still be worth mentioning to G-R for his purposes: to argue that numerical measurement of uncertainty is impossible because "there aren't enough real numbers" to correspond to lexicographic ordering. But that isn't true--assuming that set of alternatives is countably infinite--and anyway, there are other, true and strong reasons why numerical measurement of all uncertainties is impossible, even though tradeoffs between credibitility and expected value take place. G-R didn't notice that paying attention to credibility, whether in lexicographic manner or not, would conflict with Savage axioms, thus PREVENT MEASURING SUBJECTIVE PROBABILITIES IN THIS PARTICULAR MANNER: OR, CONTRADICT ASSERTION THAT PEOPLE BEHAVE "AS IF" THEY ASSIGNED PROBABILITIES, IN THIS PARTICULAR SENSE OF "AS IF."

27. Problem of "semantic information": a) subjective probabilities are involved, not objective; hence, (b) people's subjective probabilities will differ; (c) the effect of a given message on different people's subjective probabilities (e.g., assuming their a priori probs to be equivalent) will differ; (d) hence, the "information content" of the message will differ for different people, can't be defined purely in terms of the message set; (e) problem of ambiguity; in certain states of ignorance--in which messages may carry most information, or most valuable information--people may not act "as if" they assigned a priori probls at all; this will affect problem of measuring "quantity" of information even for a given person on given occasion.

28. A vn-m payoff function, (Marschak, Notes on Economics of Information, p. 85) can be regarded as "that function of the action and the state of the world, whose expected value is being maximized by the decision-maker." Its existence follows from the utility axioms.

Marschak says that "payoff function" and "criterion function" are the same thing; this implies that criterion function is so defined as to reflect risk-preferences, serving as a von Neumann-Morgenstern utility function. Perhaps this should be true, but it often isn't; often the members of a committee will agree on a criterion function to express their common goals, payoffs; yet typically the von Neumann-Morgenstern utilities of the different members will differ, though their ordering of outcomes is the same and corresponds to their criterion function.

Anyway, Savage regards an action as a function mapping states of the world onto otherstates of the world, events onto events. From this point of view, the "states of the world" which impinge on the action, the one which "obtains" being uncertain, can themselves be evaluated by the payoff function. (Note, however, that events with the same value can have very different impact upon a given action). So the action is a function mapping (specific) states with a given value onto states with a different value. We could speak of an event's being "good" or "bad" in the absence of action by the subject, or prior to a given action by him; and this is an evaluation which is distinct from the answer to the question, how "good" are the states of the world which individual available actions will map this state onto?

29. (see Marschak, Remarks on Economics of Info, p. 90):(88):

In one-armed bandit experiments, assumption that you should choose the action with higher p assumes that the advantage of "picking" a Red light is equivalent to the advantage of picking a Green light, regardless of their probabilities.

But in a given test, subjects revealed (verbally) that these two advantages came to be unequal; they actually preferred to pick the light which was more infrequent.